
Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing

A.S.Krayev, T.V.Markusheva, D.A.Kramerov, A.P.Ryskov, K.G.Skryabin, A.A.Bayev and G.P.Georgiev

Institute of Molecular Biology, Vavilov Street 32, USSR Academy of Sciences, 117334, Moscow, USSR

Received 26 October 1982; Accepted 9 November 1982

ABSTRACT

Mouse genome contains two major families of short interspersed repeats in more than 10^5 copies scattered throughout the whole genome. They are referred to as B1 and B2 sequences since they were first isolated from the genome library by means of a dsRNA-B probe /1/. In this work, two copies of the B2 family were sequenced and compared with the previously sequenced B1 repeat /2/. A B2 ubiquitous repeat is ca. 190 bp long. The members of the family deviate in 3-5% of nucleotides from the consensus sequence. B2 contains regions of homology to the RNA polymerase III split promoter and to 4.5S snRNA I. Both B1 and B2 contain regions which resemble junctions between exons and introns. In contrast to B1, B2 does not contain apparent homologies to papova viral replication origins and a human Alu sequence. One side of the B2 repeat is represented by a very AT-rich sequence (ca. 30 bp long) followed with an oligo(dA) stretch 10-15 nucleotides long. This region of the repeat is the most variable one. The whole unit is flanked with 15-16 bp direct repeats different in sequenced copies of B2. The same is true of some copies of the B1 family. The properties of B1 and B2 repeats suggest that they may represent a novel class of transposon-like elements in eukaryotic genome. A possible role of B-type repeats in genome reorganization, DNA replication and pre-mRNA processing is discussed.

INTRODUCTION

Although short interspersed repetitive sequences in the eukaryotic genome have been known for over a decade, the structural studies with these sequences became possible only after the development of DNA cloning and sequencing techniques. Many interspersed repeats seem to be transcribed and some of them give rise to the so-called dsRNA-B which generates hairpin-like structures in heterogeneous nuclear RNA /3, 4/. We have previously cloned the members of two such families designated as B1 and B2, which constitute a half and a quarter, re-

spectively of the entire dsRNA-B fraction /1/. Each of them is present in 50 to 100 thousand copies in the mouse genome, thus occurring in almost every cloned mouse DNA fragment 10-15 kb long. Only a minor fraction of these sequences was shown to exist as inverted repeats giving rise to dsRNA while most of them occur in the genome in the form of solitary copies interspersed with single-copy DNA sequences. B-transcripts constitute as much as 2% of bulk hnRNA.

The sequence of a B1 unit has been described in detail earlier /2, 5/. It is 130 bp long and contain regions of homology to junctions between exons and introns and to papova viral replication origins.

An ubiquitous repetitive sequence of the human genome, the so-called Alu-family repeat, has also been sequenced /6-9/. It possesses a number of similarities to a B1 unit and was suggested to originate from partial duplication of B1.

In the present paper, we describe the sequence of another ubiquitous repeat, B2. The structures of B1 and B2 sequences are compared to each other as well as to other sequences of the eukaryotic genome. The existence of a sequence resembling the RNA polymerase III consensus promoter in both B1 and B2 as well as the occurrence of short direct repeats at their flanks allow one to speculate that they may represent transposable elements.

METHODS

The previously characterized clones Mm14 and Mm61 /1/ were used for sequencing the two members of the B2 family. Bacterial cells were grown and recombinant plasmids were isolated as described earlier /1/. Mouse DNA insertions were separated from pBR322 DNA by ultracentrifugation in a sucrose density gradient (Mm14) or by agarose gel electrophoresis (Mm61).

Physical mapping, DNA blot hybridization and DNA sequencing methods have been described in the previous paper /2/.

RESULTS AND DISCUSSION

Mapping and sequencing of a B2 unit

Two clones carrying B2 sequences among those described

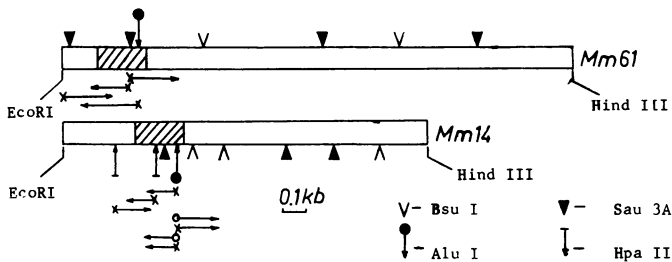


Fig. 1. A physical map of clones Mm61 and Mm14 and a strategy used to sequence areas of hybridization to dsRNA-B probe.

Arrows indicate individual sequences read starting from 5'-end labeled (x) and 3'-end labeled (o) restriction endonuclease cleavage sites.

earlier /1/ were selected for analysis. In both of them, mouse DNA was inserted into pHR322 between the EcoRI and HindIII cleavage sites. The sizes of insertions were 1.6 (Mm14) and 2.5 kb (Mm61). Fig. 1 shows physical maps of these insertions. Subfragments hybridizing to ^{125}I / dsRNA-B from mouse Ehrlich carcinoma cells are also indicated. In hybridization experiments with total mouse ^{32}P / DNA, we found that no other subfragments contained any repetitive DNA sequences. Subfragments corresponding to B2 repeats were sequenced starting from the sites indicated in Fig. 1.

Fig. 2 shows the structure of two copies of B2 sequences and of their flanking sequences as well. The borders of the repeat were determined by the disappearance of homology between the two members of the family. The length of B2 units determined as the distance between two direct repeats (see below) was found to be 190 and 195 base pairs. The two copies differed by ca. 7% of nucleotide substitutions scattered throughout the whole repeat. The region at the 3'-end of the strand shown in Fig. 2 was more variable. This region consisted of a very AT-rich sequence ca. 30 bp long and was terminated with an oligo(dA) stretch of a variable length (from 10 to 15 nucleotides).

The repetitive unit is flanked with short perfect direct repeats. Their lengths are 15 bp for Mm61, and 16 bp for Mm61.

```

Mm 6I  aattcaagagacagagataatgaaaataaccaaactaaaaattagaaattatcccaaaaaataaat
Mm 6I  aaaatggaaatcctcatcaatggaattaatcaagtggaagataaaacactaggacttgaagacaaga
      Mm 14    ...ggagctcactcaggggcagtgtagacact
Mm 6I  tagataaattagaacattcttatcataacataaatgtttaatttaagatgttaaaacatGGGCTGGT
Mm 14  tgaccaatcttgagtatagtttatattaaaatagtgacttcaagaaaatattaagtcggGGGCTGGA
Mm 6I  GAGATGGCTCAGTGGGTAAAGAGCACCCGACTGCTCTTCCGAAGGTCCGAAGTTCAAATCCCAGCAAC
Mm 14  GAGATGGCTCAGCGGTTAAGAGCAC TGACTGCTCTTCCGAAGGTCTGAGTTCAATTCCCGGCAAC
Mm 6I  CACATGGTGGCTCACAACCATCCGTAACGAGATCTGACGCCCTCTTCTGGAGTGTCTGAAGACAGCT
Mm 14  CACATGGTGGCTCACAACCATCCGTAATGAGATCTGGTGCCCTCTTCTGGAGTGTCTGAGGACAGCT
Mm 6I  ACAGTGTACTTACATATAATAATAATAATAATAATAATCTTTAAAAAAAAAAaagatgttaaaac
Mm 14  ACAGTGTACTTACATATAATAATAATAATCAACAAA  TCTTTAAAAAAAAAAAAAGAAagaaa
Mm 6I  ataacgtagatgaaataacaagaagcacaaaaaaacttgagagatttacaagacattataaaaagt
Mm 14  atattaagtcatagcatttgtggtagacaaaattggccaaatcaggaagatgacagtagatgatatt
Mm 6I  ataaaagtatgaaccttatacatataagaagaatttcaagataaaacataagacacatttcaacg
Mm 14  gggaaaacacagactgaagtgcagaaacttcagtgaaagttaaaaggttctaacgaagaaagtgcta
Mm 6I  aaattatgagaacataggccacaaaattatggtaattttcccaa...
Mm 14  tgtggcagaggcaaagggcattgtaaggtggcagtcgcacccga...

```

Fig. 2. DNA sequences from clones Mm61 and Mm14 overlapping areas of hybridization to dsRNA-B probe. Homologous sequences are given in uppercase letters. Short direct repeats flanking homologous regions are underlined. The region of B2 resembling the putative termination signal (TTCTTT) of tRNA genes /15/ is underlined twice.

The direct repeats present in the two clones are quite different from each other.

We have sequenced from 50 to 200 bp in both directions from the direct repeats. It seems that the flanking sequences for both clones are slightly enriched with AT-pairs, but no peculiar features similar for both copies could be detected.

Analysis of the B2 sequence

Since the sequence of B1 unit was reported, several papers dealing with the sequencing of rodent repeats have appeared. Oshima *et al.* /10/ studied mouse genes encoding snRNA U6 and found a sequence bearing some homology to another snRNA in the neighbourhood of the snRNA pseudogene. These authors did not

	GGGGCTGGAGAGATGGCTC AGTGGTTAAGAGCACCTGACTGCTCTTCCGAAGGTCC GA GTTCAATTC														
Mm I4	T				C			-						T	
Mm 6I			T			G		C						A	A
Mm 34	C	C						T				AG	-		
RGH I								C				A		T	
RGH 2					C		G	C				AG		AT	
CHO 49	A				GA-			- CA				AG		T	
CHO 250	A				AGA-			-			A	AG		T	
CCAGCAACCACATGGTGGCTCACAAACCATCCGTAATGAGATCTGATGCCCTCTTCTGGAGTGTCT-----															
Mm I4		G										G			
Mm 6I								C				C			
Mm 34			T					T		G				T	G
RGH I								T		A				T	
RGH 2								T		A				T	A
CHO 49								TA				G		T	CAGATATA
CHO 250			T					T				G		T	CAGATATA
----GAAGACAGCTACAGTGTACTTACATATAATAAATAAATAAA(A) ₀₋₈ TC(T) ₂₋₅ (A) ₇₋₁₄															
Mm I4			G									C			
Mm 6I															
Mm 34		TG		A		G		T	C	TG	A	-	AT	A	TT-----
RGH I										T	--		C		
RGH 2										T					
CHO 49		TATG		-	AA----	TG-		T	C						
CHO 250		CATG		-	AA----	TG-		T	C						

Fig. 3. A mouse consensus sequence of the B2 repeating unit and a comparison of it to the individual copies sequenced in rodents: a copy from mouse /10/ is designated Mm34, two copies from rat are designated RGH1 and 2 /11/, and two copies from Chinese hamster are designated CHO 49 and 250 / /. Gaps have been introduced into the consensus sequence to obtain a best fit to individual copies, whereas nucleotides missing in those copies in relation to the consensus sequence are shown as (-). A variable A-rich sequence has been abbreviated to save space.

know that, in the sequenced region, a repetitive element was in fact located immediately flanking the U6 pseudogene. This becomes clear as their sequences are compared with that of B2 (see Fig. 3). The U6 pseudogene is flanked with B1 and B2 sequences. Thus, from Oshima *et al.* /10/, we have the sequence of the third member in the B2 family which was used for constructing the consensus mouse B2 sequence (Fig. 3). The individual members of the family sequenced by us deviated from the consensus one in 4% of the substitutions. The B2 unit sequenced by Oshima *et al.* /10/ differed more from the consensus one.

Similar sequences have been found in the genome of other rodents. Page *et al.* /11/ detected and sequenced two repeats

from the second intron of the rat somatotrophic hormone gene. Its sequence differed from the consensus mouse B2 only slightly more (7% of substitutions) than did the individual mouse sequences (Fig. 3).

Two representatives of the so-called Alu-equivalent type II family /12/ of the hamster genome also appear to be close homologs of B2. In fact, they are not homologous to human Alu-sequences, in contrast to the B1 sequence. Again, the difference between the hamster B2 sequence and the mouse one do not greatly exceed the deviations between B2 copies within the mouse genome (16% and 4% deviations, respectively) (Fig. 3).

Thus, it is possible to construct a consensus rodent B2 sequence on the basis of seven individual members. Base substitutions seem to be distributed nonrandomly within B2. Some regions are conserved while, in others, substitutions are rather abundant (Fig. 4). The right-hand part of the B2 repeat, in particular, the oligo(dA) block, is the most variable.

Further analysis of the consensus B2 structure reveals the presence of homologies with some signal elements of the genome. First, there is a ca. 85% homology to 4.5S small nuclear RNA I (Fig. 4) within two regions which are located from the 3rd to 32nd nucleotide and from the 58th to 79th nucleotide.

The function of 4.5S snRNA I is obscure. Just as for some other snRNAs, it is thought to be transcribed by RNA polymerase III. Within the B2 sequence, one can also detect two segments very similar to the putative RNA polymerase III split promoter sequence deduced by Fowlkes & Shenk /14/ and Galli et al. /15/ (Fig. 4). Moreover, it is possible to find some homology between the B2 repeat and consensus exon-intron and intron-exon junctions (Fig. 4). The homology for B2, however, is less evident than in the case of B1 /2/ since it is interrupted by an insertion of the CCT block. However, this insertion is variable and sometimes consists of only one nucleotide (Fig. 3). It is possible that some B2 sequences do not contain an insertion at all.

Comparison of B1 and B2 sequences

B1 and B2 repeats apparently have different nucleotide sequences. This was expected from the absence of cross-hybrid-

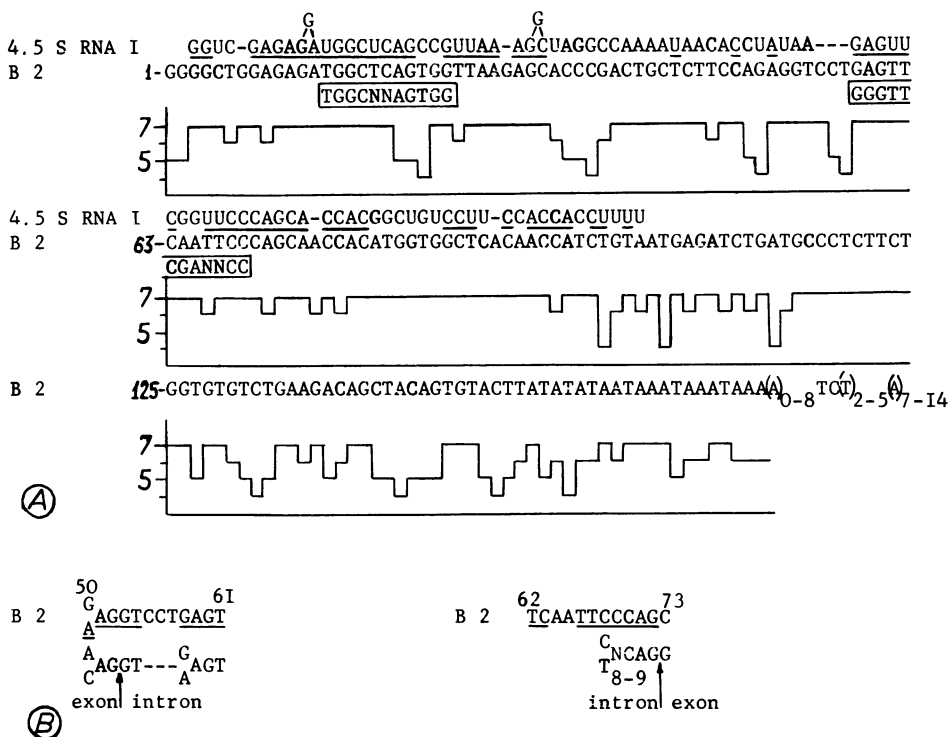


Fig. 4A. A comparison of the rodent B2 consensus sequence to the mouse 4.5S RNAI sequence /13/; gaps introduced in the latter are shown as (-). A variable A-rich sequence has been abbreviated to save space. Putative consensus RNA pol III promoter sequences /14, 15/ are boxed. A histogram displaying variation of seven individual B2 copies (see Fig. 3) is shown below the sequences; the number of sequences that coincide with the consensus one at the particular position (abscissa) are indicated on the ordinata.

B. A comparison of the regions from the B2 consensus sequence that shows homology to the exon-intron and intron-exon junction consensus sequences /16/.

zation even under relaxed conditions /1/, and becomes obvious if one compares their primary structures.

Yet, some rather short areas of homology exist between B1 and B2. The both contain two segments constituting together a structure similar to the putative RNA polymerase III promoter (Figs. 4 and 5). Regions corresponding to the second promoter segment in B1 and B2 are homologous. Fig. 5 shows three additional B2 segments (20, 31 and 13 bp long) which are 70, 71

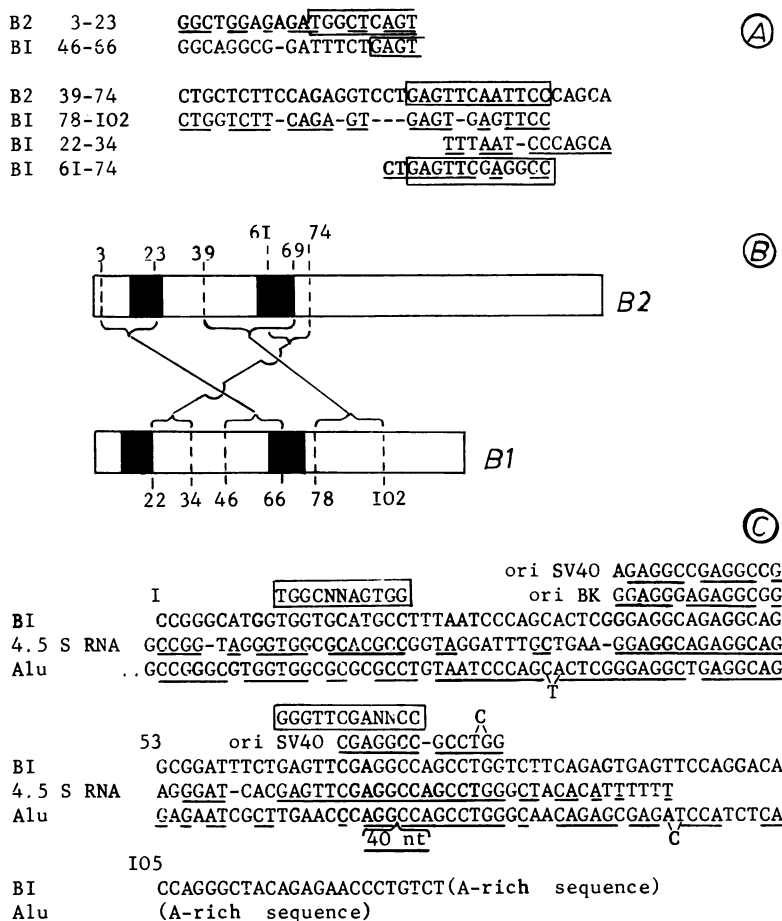


Fig. 5A. Regions of cross-homology within B1 and B2 consensus sequences. Numbers indicate nucleotide positions starting from the 5'-end of each sequence (see Fig. 4A and Fig. 5C). Positions that exactly coincide are underlined and sequences that are homologous to a putative consensus RNA pol III promoter are boxed /14, 15/. Gaps are shown as (-).

B. A scheme depicting the positions of the above homologies within B1 and B2 units. Sequences corresponding to a putative RNA pol III promoter are shown as filled boxes.

C. A comparison of the B1 consensus sequence to the mouse 4.5S RNA sequence /19/, to human Alu sequence /9/, to parts of papova viral replication origin /17, 18/, and the putative consensus RNA pol III promoter sequence (boxed). The nucleotides showing exact coincidence to B1 sequence are underlined. A gap of 40 nucleotides has been introduced into Alu sequence to obtain a proper match to the B1 sequence.

and 84% homologous to certain areas in B1. These three segments, however, are differently located in B1 and B2. For example, one of them is located at the very beginning of B2 but in the middle of B1. The significance of the latter homologies remains obscure.

In addition to the above mentioned homologies to exon-intron and intron-exon junctions /2/ and to the RNA polymerase III promoter, B1 contains two areas of homology to the papova viral replication origin (one to the center of palindrome located in the replication origin and another to a branch of the palindrome). Also long regions were observed homologous to 4.5S small RNA (which is different from 4.5S snRNA I and is found to be associated with nuclear and cytoplasmic poly(A)⁺RNAs). This small RNA has been detected so far only in rodents /19/. The B1 sequence (in contrast to B2) is closely related to the human Alu sequence, the left-hand 5'-segment 55 nucleotides long being by 82% homologous (Fig. 5C). Thus, almost all but the last 30 bp of B1 contain superimposed signal sequences. However, in spite of the above differences, B1 and B2 repeats have a number of similarities, especially in the general plan of organization. What the B1 and B2 sequences have in common is the existence of homologies to the consensus RNA polymerase III promoter and to certain small RNAs (although different for B1 and B2). The two parts of the split RNA polymerase III promoter are located at the same distance from the 5'-end in both repeats. At the same flank, the both sequences contain very A-rich regions; in the case of B2, it seems to be a part of the repetitive unit. In contrast, the A-rich segment of B1 is so variable in length and sequence that it can hardly be considered as a part of the repeat. Human Alu sequence is similarly flanked with A-rich regions, but these are less prominent.

Finally, several copies of both B1 (Alu) and B2 sequences (together with the A-rich segment) are found to be flanked with short direct repeats. Their length varies from 10 to 20 nucleotides /2, 7, 8, 11, 12, 20/. Each copy has its own direct repeats different from those flanking another copy.

All these similarities suggest that the B1 and B2 ubiquitous repeats have a similar nature.

B1 and B2 sequences as a novel class of eukaryotic transposable elements

It was suggested recently that short interspersed repeats represent a class of transposable elements in eukaryotic genome transcribed by RNA polymerase III /21/. According to this hypothesis, the primary transcript serves as a template for reverse transcriptase, and the resultant double-stranded DNA is inserted into a new position in the genome at the sites of occasional chromosomal cleavages. The above mentioned common structural properties of the B1 and B2 sequences support this hypothesis. Firstly, there is the existence of short direct repeats flanking the entire repetitive unit including the A-rich sequence. The formation of direct repeats nearly always accompanies an insertion of transposable elements in both prokaryotes and eukaryotes, including mdg elements and retroviral proviruses /5, 22/. Secondly, the hypothesis is supported by the presence of RNA polymerase III promoter sites in B1 and B2 sequences. Also, within the A-rich element of a B2 repeat, one can find a sequence resembling the RNA polymerase III terminator signal of tRNA genes /15/ (Fig. 2, underlined). Thus, the repeats may serve as adequate templates for RNA polymerase III. This was demonstrated at least in vitro for cloned human Alu-sequences /23/ and hamster B2-like sequences /12/. The 5'-end of an in vitro transcript exactly coincided with the beginning of the repeat while the 3'-end was located usually outside of the latter. Termination took place at different points where oligo(T) sequences were present. As a result, several discrete RNAs were formed up to 600-800 nucleotides long. However, it is not clear whether similar RNAs are synthesized in vivo.

The recent finding of small poly(A)⁺ RNAs containing B1 and B2 sequences, which appear to be likely candidates for postulated transposition intermediates, lends some support to the idea /24/. These RNAs with lengths of between 150 to 400 nucleotides are found in the polysome-free cytoplasm and in nuclear fractions, B2 containing RNA being much more abundant. The length of a poly(A) tail varies considerably, generating the size heterogeneity of the RNA.

However, it still has to be proved that these small RNAs

are true RNA polymerase III transcripts rather than excision products from hnRNA formed during processing.

It will be interesting, though apparently very difficult by current means, to establish whether transcription by solely RNA-polymerase III is merely a prerequisite for a B-copy transposition, or whether it is necessary for the realization of some other kind of information, which a B-repeat contains.

Relevant to this idea is the question of how many copies of B repeats are functioning (e.g. transcribed) "genes", and how many of them are in this respect "pseudogenes". Some, though speculative, evidence for the idea that only a few of B-copies do function as RNA-polymerase III templates producing analogs of snRNA comes from comparing the sequences flanking B1, B2 and U6 pseudogenes /10/. The occurrence of A-rich stretches at the 3'-end of all three repetitive elements may well be a coincidence, but otherwise may reflect the operation of a common "switching" transposition mechanism, similar to that suggested in /21/, and involving a poly(A) RNA intermediate.

We believe, that the "transposon hypothesis" accounts for many important properties of the repeats of the B-type.

The long transposable elements or the so-called mobile dispersed genes (mdg) seem to be widespread in eukaryotic genomes /5/. However, these elements are much larger (5-10 kb) and are transcribed by RNA polymerase II. Their insertion also creates a duplication of the target sequence, but only 4-5 bp long. Also they are less abundant in the genome than B-type repeats. Finally, the mdg elements are less conservative. The mdg sequences differ in the genomes of closely related species. At the same time B1-like repeats have been reported for primates /9, 25/, rodents /2, 3, 20, 26/, and probably chicken /27/, while B2 exists in all rodent species studied to date (see Fig. 3).

Do B1 and B2 sequences possess any biological functions?

The intriguing question is whether short interspersed repeats are some type of a selfish DNA capable of transposition or represent useful information for the genome, or both. Considering the homology between B-type repeats and certain snRNAs

it has already been suggested that the former are truncated pseudogenes for 4.5S RNAs. This is unlikely as B-type repeats are longer than homologous snRNAs, their "extra sequences" being quite conservative. Yet, B-type repeats and some small RNA genes may be in some kind of phylogenetic relationship.

The conservative character of ubiquitous repeats and the presence of signal sequences within them suggest that they can play some functional role. Their homology to snRNA, for example, may be explained by similar functions. Though the role of most snRNAs remains unknown, speculations about their involvement in splicing are common and have found strong experimental support, at least in the case of snRNA U1 /28/.

A possible mechanism of such involvement is the formation of duplexes with intron ends resulting in alignment of the adjacent exons to each other. It is likely that other small RNAs, for example, 4.5S snRNA I (which is homologous to B2) and 4.5S and 7S RNA (homologous to B1) and also B1⁺ and B2⁺ RNAs themselves are involved in RNA processing in a similar fashion. Otherwise, B1 and B2 sequences may serve as "substrates" in RNA splicing, being recognized in pre-mRNA by certain snRNAs.

It is very likely that the major part of B1 and B2 transcripts is included in hnRNA transcribed by RNA polymerase II /1, 22/. A significant fraction of 4.5S RNA is associated in vivo with hnRNA, probably forming duplexes with B1 sequences. The presence of segments resembling exon-intron junctions in B1 and B2 repeats may be attributed to the B1 involvement in splicing. Similar signal sequences were found in quite different repetitive elements of sea urchin /29/. Splicing is known to proceed stepwise if "junction sequences" are present inside the intron /30/. Some authors suggested that gene expression can be modulated at the level of processing and splicing /31/. Thus, the existence of a B1 repeat within an intron may create multistep splicing and, in this way, influence gene expression.

Some sequences of the B-type were shown to survive the splicing and appear in mRNA. This was originally shown by mRNA-dsRNA hybridization /4/ and recently confirmed by mRNA-B1 or mRNA-B2 sequence hybridization /24/. At least some of B1 and B2 sequences are located at the 3'-end of poly(A)⁺ RNA /32/.

Possibly, some signal sequences within B-units, e.g. AATAAA boxes are used as polyadenylation and termination signals for large RNA polymerase II transcripts.

Another possible role of the repeats of the B-type is to serve as sites for the initiation of replication. The hypothesis is based on the apparent homology between B1 sequences and the replication origins of papova viruses. Segments adjacent to the replication origin were reported to be enriched with B1 sequences /5/. Alu-sequences were found once in each ribosomal DNA repeating unit in human DNA /33/. However, no direct evidence for their function as origins of replication has been obtained so far.

Other hypotheses suggesting how B1 and B2 sequences may be involved in the regulation of gene expression are possible. For example, they may serve as elements of a modulator type influencing the transcription rate, as structural elements determining RNA interaction with the nuclear skeleton during RNA transport, etc. It also remains to be elucidated what particular type of regulation, if any, operates through transposition. Anyhow, if a B-type sequences can influence some steps of gene expression, the transposition of B-copies into new sites of the genome could have generated many cryptic differences among individuals, and, in this way, accelerated the evolutionary process. This can be illustrated by a recent sequencing study of goat globin genes /34/, three of which, being developmentally regulated, are apparently identical, except for short inserted (deleted) repeated sequences within introns. At least by analogy, this demonstrates how the integration of a B-like element could have influenced gene regulation, producing cryptic differences between individuals subject to further selection.

ACKNOWLEDGEMENT

The authors are greatly indebted to Dr. A. Janulaitis and his colleagues from the Institute of Applied Enzymology (Vilnius) for numerous gifts of many restriction nucleases and T4 polynucleotide kinase.

REFERENCES

- 1 Kramerov, D.A., Grigoryan, A.A., Ryskov, A.P., Georgiev, G.P. (1979) Nucl. Acids Res. 6, 697-713.
- 2 Krayev, A.S., Kramerov, D.A., Skryabin, K.G., Ryskov, A.P., Bayev, A.A., Georgiev, G.P. (1980) Nucl. Acids Res. 8, 1201-1215.
- 3 Ryskov, A.P., Saunders, G.F., Farashyan, V.R., Georgiev, G.P. (1973) Biochim. Biophys. Acta 312, 152-164.
- 4 Kramerov, D.A., Ryskov, A.P., Georgiev, G.P. (1977) Biochim. Biophys. Acta 475, 461-475.
- 5 Georgiev, G.P., Ilyin, Y.V., Chmeliauskaite, V.G., Ryskov, A.P., Kramerov, D.A., Skryabin, K.G., Krayev, A.A., Lukanidin, E.M., Grigoryan, M.S. (1981) Cold Spring Harbor Symp. Quant. Biol. 45, 641-654.
- 6 Rubin, C.M., Houck, C.M., Deininger, P.L., Friedman, T., Schmid, C.W. (1980) Nature 284, 372-374.
- 7 Bell, G.I., Pictet, R., Rutter, W.J. (1980) Nucl. Acids Res. 8, 4091-4109.
- 8 Pan, J., Elder, J.T., Duncan, C.H., Weissman, S.M. (1981) Nucl. Acids Res. 9, 1151-1170.
- 9 Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T., Schmid, C.W. (1981) J. Mol. Biol. 151, 17-33.
- 10 Oshima, Y., Okada, N., Tani, T., Itoh, Y., Itoh, M. (1981) Nucl. Acids Res. 9, 5145-5156.
- 11 Page, G.S., Smith, S., Goodman, H. (1981) Nucl. Acids Res. 9, 2087-2104.
- 12 Haynes, S.R., Jelinek, W.R. (1981) Proc. Natl. Acad. Sci. USA, 78, 6130-6134.
- 13 Ro-Choi, T.S., Reddy, R., Henning, D., Takano, T., Takano, T., Taylor, C.W., Busch, H. (1972) J. Biol. Chem. 247, 3205-3222.
- 14 Fowlkes, D.M., Shenk, T. (1980) Cell 22, 405-414.
- 15 Galli, G., Hofstetter, H., Birnstiel, M.L. (1981) Nature 294, 626-631.
- 16 Mount, S.M. (1982) Nucl. Acids Res. 10, 459-472.
- 17 Yang, R.C.A., Wu, R. (1979) Science 209, 456-462.
- 18 Reddy, V., Thimmappaya, B., Dhar, R., Subramanian, K., Zain, S., Pan, J., Celma, M., Weissman, S. (1978) Science 200, 494-502.
- 19 Harada, F., Koto, N. (1980) Nucl. Acids Res. 8, 1273-1285.
- 20 Haynes, S.R., Toomey, T.P., Leinwand, L., Jelinek, W.R. (1981) Mol. Cell. Biol. 1, 573-583.
- 21 Jagadeeswaran, P., Forget, B.G., Weissman, S.M. (1981) Cell 26, 141-142.
- 22 Temin, H.M. (1981) Cell 27, 1-3.
- 23 Duncan, C.H., Jagadeeswaran, P., Wang, R.R.C., Weissman, S.M. (1981) Gene 13, 185-196.
- 24 Kramerov, D.A., Lekakch, I.V., Samarina, O.P., Ryskov, A.P. (1982) Nucl. Acids Res., this issue.
- 25 Grimaldi, G., Queen, C., Singer, M.F. (1981) Nucl. Acids Res. 9, 5553-5568.
- 26 Kramerov, D.A., Lomidze, N.V. (1980) Dokl. Akad. Nauk USSR 225, 1005-1009.
- 27 Stampf, W.E., Kristo, P., Tsai, M.-J., O'Malley, W. (1981) Nucl. Acids Res. 9, 5383-5397.

- 28 Lerner, M.R., Steitz, J.A. (1981) Cell 25, 298-300.
- 29 Posakony, J.W., Scheller, R.H., Anderson, D.M., Britten, R.J., Davidson, E.H. (1981) J. Mol. Biol. 149, 41-67.
- 30 Avvedimento, V.E., Vogeli, G., Yamada, Y., Maizel, J.V., Pastan, I., de Crombrughe, B. (1980) Cell 21, 689-696.
- 31 Davidson, E.H., Britten, R.J. (1979) Science 204, 1052-1059.
- 32 Ryskov, A.P., Ivanov, P.I., Kramerov, D.A., Georgiev, G.P. (1982) Dokl. Akad. Nauk USSR 263, 997-1000.
- 33 Higuchi, R., Strang, H.D., Browne, J.K., Martin, M.O., Huot, M., Lepeles, J., Salser, W. (1981) Gene 15, 177-186.
- 34 Schon, E.A., Cleary, M.L., Haynes, J.R., Lingrel, J.B. (1981) Cell 27, 359-369.